



Building a Statewide Address Location Database Current Status, Issues, Plans

MassGIS / Wareham, MA
February 4, 2013

Information Technology Division
Executive Office for Administration and Finance



Workshop agenda

In partnership with State 911, we are creating a statewide map of address locations

This map will be more detailed and more complete than any currently available map

In this workshop, we'll describe:

- Current and future 911 use of GIS
- The need for standards and a statewide approach
- What's been produced so far
- What we will complete this year
- What will remain to be done

2

We are here because we, MassGIS, in partnership with state 911, are working on a project to develop a statewide map of address locations and we want you all to know what we are doing and to offer you the opportunity to get involved with this project going forward. State 911 is our partner in this project because what we are doing will support the new 911 systems to be deployed over the next few years, the so-called Next Generation 9-1-1 systems. That's why we've invited public safety officials to these workshops. But we believe that the GIS data products that we'll be discussing are very valuable for other municipal departments, so we've also invited municipal GIS staff and other GIS users as well. Our goal is to engage with anyone and everyone who is interested in GIS and addressing. The benefits from this project will be very broad. In fact, we've already created millions of dollars worth of GIS data for municipalities to use and in the process we've saved local governments lots of money that they might otherwise have spent.

There are a number of reasons we want you all to know what we are doing:

You may be interested in creating or using products similar to those that we are creating - by knowing what we have done you can avoid re-doing the same work

Or conversely, you may have already worked on some aspect of this and maybe we can avoid re-doing any work that you've already done. All the mapping that we are creating is being shared with regional and local government. I'd like to point out that's not always

the case with 911 address information, for instance in NH the data are being very closely held and not shared. We believe that taking an open approach means involving more stakeholders and ultimately creating a better product.

you may be wondering is this a typical state agency “top-down” approach which will impose a solution that isn’t right for every municipality. Its true that we’ve spent a lot of time figuring out the best way to do this on a statewide scale... As you can imagine, it involves some very large databases, millions of address records to be reconciled, a lot of automation, computer programming to do things quickly and efficiently and many thousands of address locations to be verified in the field. But we’ve been thinking a lot as well about how and when to involve localities in the effort. It may be that we are creating is not exactly what you want or need for all purposes, maybe not detailed enough or designed a little differently from what you already have, but we’d still like to talk about how we can share information. Since what we are creating is a statewide product which is based on national standards, it does have to be “one-size-fits-all” but on the other hand, we have listened carefully to municipal users and tried to extend the product so that it provides as much detail as anyone could want. Going forward, we need to work together to improve the product because after we’ve done whatever we can at the state level “on the desktop” there will still be many thousands of locations that need to be verified.

First, let me give you the background on this project, what we are doing and why. Then we can come back to the questions around state-local cooperation.



Workshop agenda

We want to hear from you as well:

- **Are you trying to solve the same problem at the local level?**
- **What can we learn from your efforts and experience?**
- **How are you currently managing address data?**

Most important, we want to engage you in this effort:


- **We are already working with many cities and towns**
- **A GIS-based address database is a resource for municipalities**
 - public safety, permitting, notifications, schools etc.
- **The first version will require field work to complete**
- **We will need local involvement as well to maintain it**

3

It will take about an hour to give you the background on this project, then we'll take a break and get into discussion. We want to hear about your needs and efforts at the local data, and how you are managing and using address data. What can you tell us that may be helpful in building the statewide database. For example, with dispatch, we would love to hear about what kinds of situations are problematic. Is it trailer parks, campgrounds, condo complexes? most importantly, we want to connect. I'm here with Michael Warner who is our full-time municipal outreach person. He is working with 911 data managers and public safety officials on some aspects of this which I will cover later in the presentation. In a few months, we will be ready to start the field data capture phase of this project, and he will be coordinating those efforts as well. We hope to engage with local government, with firefighters, planners, IT people or whoever, to finish the work and begin maintenance.

Current Technology

First I want to talk about current technology, how we are supporting 911 and other users of address data in GIS.



GIS and 911 requirements – current technology

- **Some 911 terminology**
 - PSAP = Public Safety Answering Point
a 911 call is routed to a PSAP where a dispatcher manages the response
 - NG-9-1-1 = Next Generation 911
standards and specifications for modernized 911 which will be GIS-driven, IP-based, and support multiple protocols
 - ESN, ESZ = Emergency Service Number, Emergency Service Zone
unique combination of PSAP and police, fire, medical responders
the ESZ is the area for a single ESN
 - MSAG = Master Street Address Guide
list which assigns an ESN to every valid, contiguous ranges for every street in a community
 - ESL = Emergency Service List
complete list of land line numbers, new addresses validated against MSAG
- **Current technology – display caller location**
 - Linear geocoding is used to plot addresses coming in with the call
 - Reverse geocoding used to show address nearest cell call

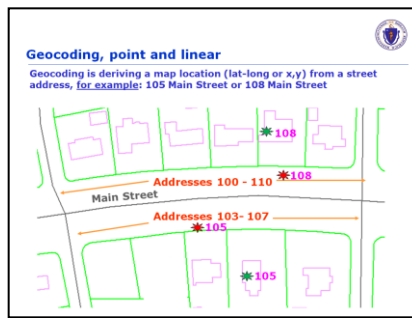
9

First, some 911 technology, since we'll be talking about what we are doing to meet their requirements. But Again, throughout this presentation, please remember that the same capabilities can be used by other departments and for many different purposes. Here are some acronyms and definitions that we should introduce.

READ the definitions....

Right now, we are now providing a linear geocoding capability for state and municipal users which is based on the NAVTEQ street map with address ranges. This dataset is being customized for 911 to display caller locations based on landline address in the PSAPS. Cell phone locations plot based on the lat-long reported by the phone. Currently, **routing** of calls is handled with a tabular lookup of the phone number. cell phone calls are handled by the state polices. one issue that we want to discuss is that the linear gecoding results and the cell phone locations that come in with the call may be very much in error which can lead to serious problems.

As you all know, Next Generation 911, is GIS-driven. That means that the accurate mapping of caller locations and the mapping of PSAP boundaries using the ESNs become **critical** because point-in-polygon overlay is used to route the call --that's 911's major responsibility, to route the call Also, in Next Gen environment you'll see much more sophisticated dispatch software, with imagery, links to other databases, automatic vehicle location and other fancy stuff, so accurate display of caller location, will be important for dispatch in that environment.



Here's an illustration of how geocoding works. You see here a street segment, a block on main street and you see that on the left and right sides you have two address ranges – 100 to 110 and 103-107. If you're trying to get to 108, and all you have is the road centerline with its address range, then you estimate the location of 108 accordingly. This is called "linear geocoding" But as you can see on the bottom, that doesn't always work very well sometimes – number 105 estimated to be in the middle of the block winds up in front of the wrong house. This example is fairly tame and, for many situations, the result is "good enough" ; however, linear geocoding can produce results that are often much worse. The reason for this error is that the address location is interpolated based on the assumption that the addresses in the range are evenly distributed along the street block. Geocoding to the parcel level, if the parcel has an address, is necessarily going to be more accurate.

Geocoding is very useful for anyone who wants to see the approximate location on a map of where they want to go, who is planning service deliveries, or who wants to look at the demographics of their client base. But I want to stress the word approximate, because we've seen errors up to several hundred feet with linear geocoding.

Current technology - Mapstar in PSAP



Shown is an address plot in Mapstar. This software runs in all PSAPs but many use Computer Aided Dispatch systems as well.

7

Linear geocoding is what displays now in the PSAPs, as I said. This is the mapstar display. CAD systems may provide more information, such as display of structures and individual address points.

Current technology - linear geocoding

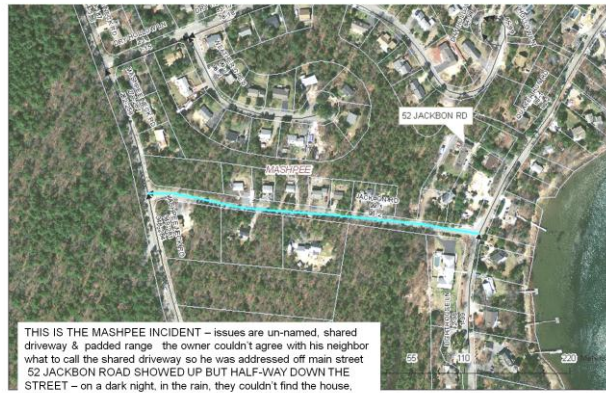


- **We provide street map for PSAPs to geocode**
 - based on product from NAVTEQ, commercial provider
 - edited extensively by MassGIS to work with Mapstar
- **We task NAVTEQ to capture missing and new streets**
 - quarterly submission of Map Requests to collect new streets and ranges with GPS-enabled vehicles
- **Since the project inception, the street map has dramatically improved**
 - 45,000 edits, 4,000 new streets
 - currently about 100-200 requests / quarter
- **Geocoding hit rates: match to MSAG**
 - September 2007 - 87% of street segments
 - October 2012 – about 99.5% (99.8% ESL)
 - 3000+ developed streets not in MSAG



The linear geocoding solution that we are providing is based on a statewide map of streets and address ranges from a commercial provider of car navigation data, NAVTEQ; we look for new or missing streets in the msag and other sources and we send them data update requests; they find and drive those new streets and with the vehicles you see here the GPS on the roof, and technicians operating several high tech data collection devices like ground based LiDAR. Since the project inception, we've updated 45,000 records and added 4,000 new streets. currently, the updates are running about 100-200 a quarter- of course we aren't seeing a lot of development right now. the geocoding hit rate is the percentage of street segments listed in the MSAG that we can match, at about 99.5% - the actual hit rate for individual addresses is higher, about 99.8%. but one interesting statistic, and one that is likely to increase, is the number of streets that should be in the MSAG and aren't because there is no-one with a landline living on them.

Current technology - why we need point addressing



THIS IS THE MASHPEE INCIDENT – issues are un-named, shared driveway & padded range the owner couldn't agree with his neighbor what to call the shared driveway so he was addressed off main street 52 JACKBON ROAD SHOWED UP BUT HALF-WAY DOWN THE STREET – on a dark night, in the rain, they couldn't find the house.

THIS IS THE MASHPEE INCIDENT quite possibly familiar to some of you. it illustrates a number of issues with mapping that are incredibly common. the caller's address was 52 Jackbon road – although the address was on a shared driveway and the town had asked the residents to name their street.. we have JACKBON ROAD MAPPED AND SO it geocoded but because the range is “padded” as 0-100 number 52 plotted HALF-WAY DOWN THE STREET. it was a dark and stormy night and the ambulance was delayed because they couldn't find the driveway entrance. There were other complicating factors, but IF WE WERE GEOCODING to point addresses there might have been a better chance of getting there in time. There are hundreds of un-named, shared driveways across the state and they are scary – look at this one.....



the house is a mile from the road, this is in dartmouth, at the end of a shared driveway and you have to wonder what signage there is at the fork and what if it was in the dark, bad weather and if someone took a wrong turn that would cost a lot of time. of course the driver and the caller and the dispatcher are all patched in, so someone presumably is giving directions, but it's still the kind of situation where you would like the dispatcher to have as much information as possible.

**Future Technology
Next Generation 911**

12



Future technology - Next Generation 911

- **National Emergency Numbering Association (NENA) is developing standard for next generation of 911 technology**
- **NG-9-1-1 will use caller locations (GIS points) and emergency service zones (GIS polygons) to route calls and assist dispatch..**
- **for routing:**
 - caller location comes from GPS in cell phone or from matching address from landline
 - emergency service zones identify PSAP
 - point-in-polygon overlay is used to assign the call
- **for response:**
 - caller location is displayed with other map layers to assist in dispatch

13

I've mentioned that next generation 911 is GIS driven – here's the system will work, a GIS point representing caller location will plot inside a GIS polygon representing the psap area and that's how the call will be routed – we call that a point-in-polygon overlay and obviously for it to work properly the location of the address has to be accurate and errors in linear geocoding could be much more serious than they are now. As we just saw, the more accurate location will also benefit dispatch, and we expect that in next gen systems the software in the psaps will include display of much more GIS data – perhaps imager, structures and other layers.



Future technology – design considerations

To support call routing and enhanced dispatch in NG, we need:

- *a location for all addresses*
- *a map of emergency service zones*

Design Considerations in Massachusetts

1. Parcel data represent the most obvious initial source for address location
2. Parcel data can be aggregated to represent emergency service zones
3. Multiple sources of address information must be conflated
4. Best to represent address locations by visible features such as buildings
5. Assignment of addresses should minimize interpretation
6. Many addresses may be associated with one location
7. Many-to-many relationships difficult to manage; use multi-points
8. Sites like campuses, large condo complexes need more address detail

34

So given the disadvantages of current technology, we're looking forward to Next Gen. As I described earlier, to make Next Gen work, we need a location for every address and we need an accurate map of emergency service zones. What I've put up here is our analysis of the problem, and it's worth spending a couple minutes on it. First, we realized that parcel data would be the most cost-effective source for the initial round of addressing, if we could match parcel data to the Emergency Service list.

we also knew that we wanted to associate addresses with actual structures visible on the ortho – ortho is a reality check and of course the more up-to-date the better. in our work with Broadband we had learned that it was a mistake to try to interpret which was the primary structure on a parcel – that's a key point. So we determined to address all structures. Of course, many addresses might be associated with any one location or even one structure, so from the beginning we thought of taking a relational approach. But from bitter experience with parcels, we wanted to avoid a many-to-many situation which is hard to manage in a GIS so we determined to use multi-point features for the address points. Finally, we planned to collect additional address detail through field work , where large sites like college campuses or trailer parks might have many structures that needed to be addressed individually by name or by number. That's the stage that we are at now, about half way through processing the parcel addresses and beginning the field work.

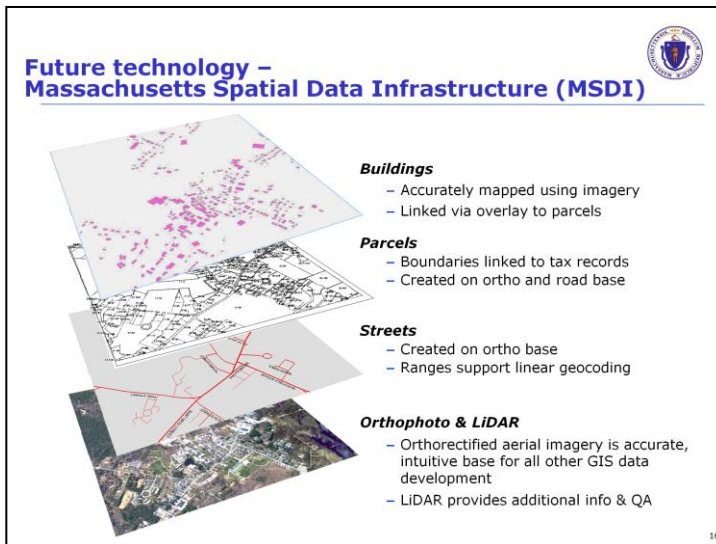


Future technology – data requirements


- **updated imagery**
 - **updated street map with ranges**
 - **standardized tax parcels**
 - **outlines of all structures**
- plus...**
- **standardized addresses from various sources**
 - **emergency service zones**

15

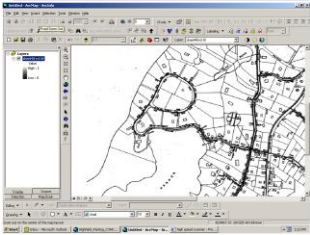
This is the list of GIS data that we know we need, with the emphasis on currentness and completeness . These are the inputs to building a next gen database – imagery and streets, which we need for current technology plus parcels, (linked to assessor property records at 99.5% match rate), structures > 150 sf and of course address lists. From the parcels we create emergency service zones as well as deriving point addresses..The good news is that this is also the foundation of a multi-purpose GIS for municipal use, just what we identified in our 2007 strategic plan for spatial data infrastructure. The bad news is that we have an unbelievable amount of work to do to pull together all this information, process it, deploy it and of course maintain it.



Here you see the GIS layers, that make up the MSDI Mass spatial data infrastructure we call this the pancake, Starting from the bottom, orthophoto is critical because it's the base on which we compile all the other information. On top of the ortho we are continuing to maintain the streets layer for geocoding purposes as I described. The streets and the ortho provide the framework for the tax parcel mapping which comes from local govt and is being aggregated and standardized at the state level. Finally, built on top of all the other layers are building outlines. All these layers are complete for Mass except the parcel layer, which is almost complete, the structures we just finished a few months ago based on spring 2011 and spring 2012 imagery.

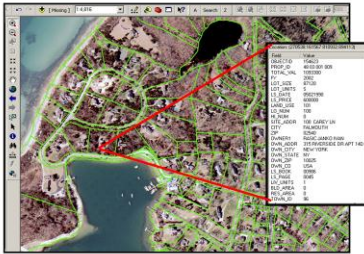


Parcel data – initial address locations



tax records provide site address, owner name, use code, building value and year built

local tax maps are brought into GIS and matched to aerial photo, assigned a unique statewide ID and linked to the tax records at 99.5% match

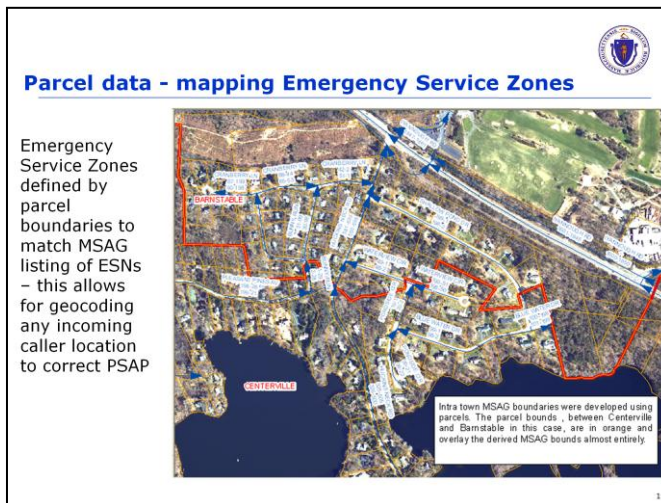


3 year - \$3M project to standardize about 2.2. million parcels; funded primarily by IT bond. Substantial economies of scale compared to individual jurisdictions, 3x or more.

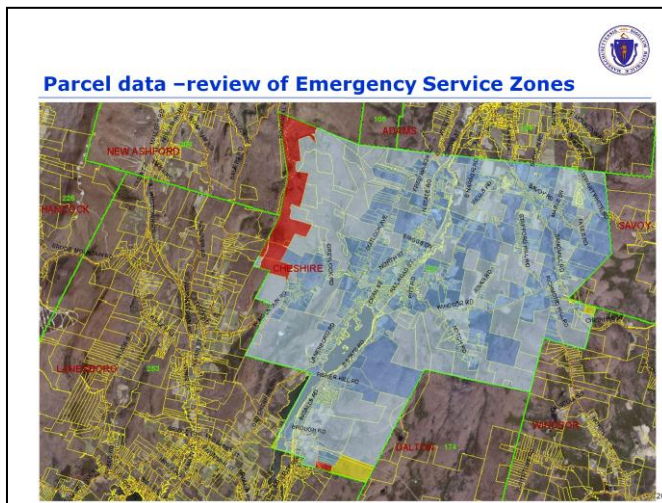
the parcel mapping has been a huge project - we are just finishing up the last of all 351 cities and towns except Boston upgrading and standardizing the local parcel data, be it paper map, CAD , poor quality or wonderful quality GIS, it has all gone into the statewide database. The key requirements are a very high match to the tax list, 99.5% and the parcel boundaries have to be accurate enough so that structures fall into the correct parcel. Unless of course they don't in reality, which they often don't. IT took us three years and three million dollars to automate 2.2 million parcels and we kept six vendors busy. We are doing the work of standardizing the parcel addresses in-house, that is the single most time-consuming step in the process, even though it's highly automated, because this is local data and we have no control over data format.



this little sequence shows how poor quality parcel data, here on the boundary between new bedford and acushnet, is being cleaned up so that it correctly registers with the imagery and with the structures.



here you see another reason why 911 needs parcels, to build the esz's which will determine how calls are routed in the next Gen software. To create these we match the parcel addresses to the MSAG ranges which list the ESN and then aggregate the parcels . this is from cape cod which is really balkanized in terms of response areas. Because of mistakes in the MSAG and what are called overlap/underlap issues, which I will illustrate on the next slide, this part of the project will have taken us three times as long as we figured and will be complete in about five months.



The highlighted area is Cheshire, a town in Western MA and the part colored red is to the west of a mountain range makes that part of town only accessible from the neighboring towns, so the ESZ boundary has to be different from the jurisdictional boundary – this is called underlap in 911 parlance. Another problem with ESZ's is that Verizon often did not correctly list the MSAG community in the MSAG, which has caused us enormous difficulty because we actually have to go look at every parcel along every towns boundary. We are seeing dozens of MSAG anomalies in rural towns and sometimes hundreds per community in more developed areas.



Parcel data - ESN boundary issues

This is a close-up of one situation on the boundary between Dracut (on the left) and Methuen. McGrath Road is in the MSAG for Dracut but not Methuen. The red is an unassigned parcel, meaning we can't determine assignment from MSAG and blue is assigned to the other community. Blue parcel at the bottom is on "North Lowell Street" in Methuen but egress is to Dracut so that has to be reviewed



21

Structures – complete statewide.

Structure outlines were completed statewide at a cost of about \$0.13 per outline for about 2.4 million structures

The centers of these building outlines, which we can automatically generate in the GIS, provide the initial set of points to be linked to the parcel address.

These data are now available on our website..




2

here's the last MSDI input – statewide mapping of structures, compiled from orthophoto. We're pretty happy with this – 2.4 million structure outlines costing about \$0.13 each from Rolta International and comparing pretty favorably to planimetric products costing much, much more. I know.. too much late night TV. The centroids of these outlines provide the points for the initial draft address location database. Every point must be linked to an address.

Address data - what is an address anyway?

- Thoroughfare address = number, street, location**
 - parts of an address must be distinct
 - location also called sub-address includes any info NOT part of number or street name like unit, building name, etc.
 - 20 | Main Street | Unit 4
- Functions of an address**
 - how you physically find a location
 - how you uniquely identify something
 - how you send mail to person or business
- Address locations can be more or less accurate**
 - unit within bldg.
 - entrance to a bldg. with signage (this could be a separate unit)
 - a building itself
 - identifier for parcel of land
 - entrance to campus with multiple bldgs.
 - estimated location using linear geocode



← Sweet spot for public safety

← this may be perfectly adequate


← or even this

← what we have now

I want to just focus for a minute on what an address is and how we use them. First of all, a lot of our work involves breaking an address that we get into its component parts, number street and then everything else, what the ESL calls “location” and the post office calls “sub-address”. For instance, the example here is 20 Main Street, unit 4 – that unit information is the sub-address. Now an address can be how you find something, it can be an identifier, it can also be how you send mail, which is really how the mailman finds something.

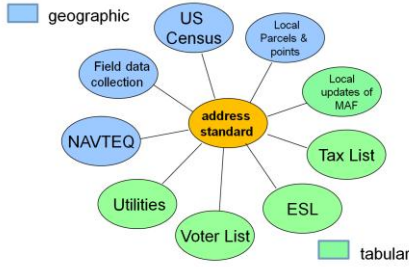
if you’re trying to find someone, their address is their front door, quite possibly a unit inside a building. if you’re the mailman, you aren’t climbing up to the fourth floor to deliver the mail – most likely the mail boxes are near the entrance to the building, which has some signage on the outside. From a first responder point of view, the door to the unit is good, but you’d settle for the building. From the perspective of a local assessor, the address might be associated with several buildings or even with a undeveloped parcel of land. The address could represent the entrance to a whole complex of buildings, like the gate you see here, you could have a single street address for an entire college campus. That gets you get into the secondary portion of the address, the building, floor and unit.

The Federal Geographic Data Committee developed a standard for addressing which NENA the national emergency numbering association adopted and which lays all this out.

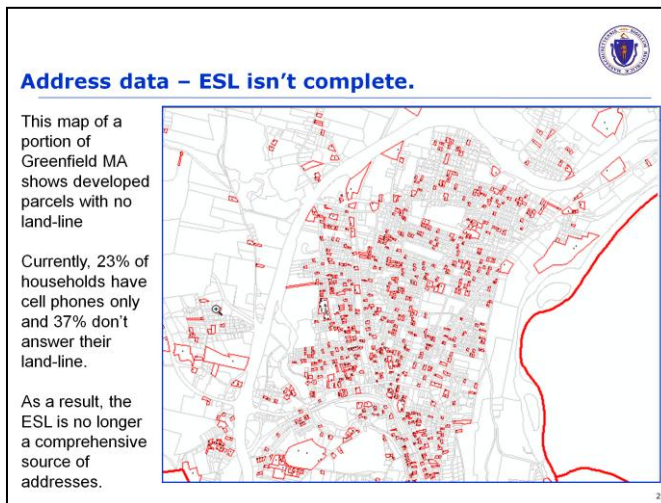


Address data - why is a standard so important?

- A standard allows for automated matching between many different address lists & mapping sources
 - Emergency Service List is no longer complete list
 - Each candidate address needs to be matched against existing list
 - *Geocoding must be exact, not fuzzy match*



We need an address standard because we have no other way of matching between different sources of address information. What I'm showing here is tabular sources in green and geography in blue. the drill is to match them. As I just mentioned more and more people are cutting the land line, the ESL is no longer a complete list. We are using the local tax lists as the primary source, but they are not real time, Besides tapping into the 351 different address assignment workflows at the local level in Mass. which we have documented but are sometimes pretty tenuous, utilities are probably the holy grail of getting new addresses. The address standard ensures that we can reliably identify new addresses vs. ones we already have, and do the outreach and/or field work to map them if we don't have that already.



I've mentioned the ESL a few times – it's the white pages (plus the unlisted numbers) and is used to lookup the address by phone number in the current technology. Here's why the ESL doesn't work for us – its because, as you all know, landlines are going away, One survey found that currently, 23% of households have no land line at all, only a cell phones. So those addresses are now missing from 911's master list. This is why we are conflating our master address list from multiple sources. There is no single comprehensive source, although electric utility customer lists will come close they won't list all rental units.



Address data standard from NENA and FGDC

Cleaning up and standardizing address information is a major challenge but NENA requires it. We first parse addresses into three parts – number, street and location (sub-address), then further parse those into the fields below.

Address Number – Prefix, Number, Suffix

- prefix "B18", "Milepost 12.2"
- number "247"
- suffix "12A", "14 1/2"
- ranges are stored in two number fields to allow for inequality testing

Each part of a street name has to conform to the specification from NENA with allowable values for street type and so on validated against a reference list.

Street Name – standard says everything fully spelled out

- pre-mod "Old North Coach Road", "Upper Hampden Road"
- pre-dir "South Main Street"
- pre-type "Avenue A"
- street-name "Broadway"
- post-type "Market Street"
- post-dir "Washington Street South"
- post-mod "Charles Street Place", "Chatham Street Extension"



Address data - Master Street Table based on standard

SOURCETYPE	FULL_STR	FULL_STR_STD
TIGER	Garden Street Ct	GARDEN STREET COURT
BASE_STREETS	GARDEN ST CT	GARDEN STREET COURT
NAVTEQ	GARDEN STREET CT	GARDEN STREET COURT
MSAG	GARDEN ST CT	GARDEN STREET COURT
WARREN_GROUP	Garden Street Court	GARDEN STREET COURT
MSAG	E Spring St RR Xng	EAST SPRING STREET RAILROAD CROSSING

MassGIS has completed a lookup of the street name in a number of statewide sources (TIGER, Warren Group, MSAG, DOT, NAVTEQ) to a standardized version. Discrepancies in type or spelling have been reconciled as well as expanding all abbreviations to conform to NENA standard.

27

TIGER is the census mapping



Address data – sub-address

The ESL, tax lists, utility customer lists, voter lists and other sources of address information will frequently include unit information. Also common are site names (landmark addressing) and in the ESL at least identification of specific buildings or even floors. This information must be standardized as well before address lists can be compared.

Sub-address (ESL "location", postal "secondary location")

- site "Bristol Community College, Administration Building"
- subsite "Jackson Athletic Complex, Field C"
- building "Mission Hill Condos, Building A", "Admissions Building, UMASS"
- rel_loc "47 Market Street, left side" "22 Maple Drive, rear"
- floor "2nd Floor, Program Office"
- unit "Unit 7A"
- own_occ "Dunkin Donuts", "Mariott Hotel" (also called landmark name)

28

Back to the address standard for one second, here's the standard as it applies to sub-address information, we've slightly expanded it. There are separate fields for each of the parts for the sub-address. This is where you include information about sites, (read examples) and locations within a site like parking lots or playing fields, this are also fields for building name, possibly business or institution name . The building, floor, unit are all fields in the NENA standard.



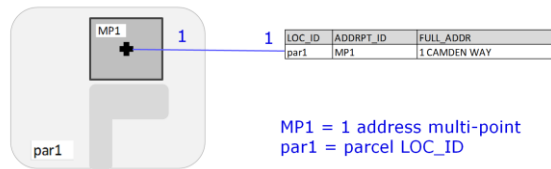
Address location data management

- Every address links to just one location
- An address location can be:
 - parcel centroid (pre-assigned address)
 - non-building point such as playing field or parking lot
 - cluster of building center points (multi-part point)
 - building center point
 - building entry
 - building interior
- Every address location must be geographically distinct and must convey "real" information about that location relative to other locations
 - no stacked points
 - no "shotgun" or arbitrarily placed points
- An single address location is often linked to more than one address
- Center of building cluster can be used for labeling

29

So how are we going to organize and manage all these datasets. Here are what you might call our principles of address location data management why we are organizing and managing our data a specific way. The most basic rule is that every address links to an address location (geography) via a shared address point ID. an address location can be a parcel center, it can be a non-building point like a parking lot, it can be a cluster of buildings grouped together into what we call a multi-point feature, it can be a single building center point, it can be a building entry door or an interior point. The key is that every address location represents a unique and distinct location –In a lot of systems where every address has its own point, a 100 condo units in a building will be represented by 100 points all stacked up but we think that's too awkward to manage, so we don't allow stacked points or another version that you see sometimes, arbitrarily scattered points. Also, we call this a relational model, because a single address point can be linked to a lot of addresses, as in an apartment building with hundreds of units. Of course with clusters of points you probably just want one label.

Simplest case – one parcel, one structure, one address



In the simplest case, we transfer the parcel address to the address point created as the centroid of the building.

I'm sparing you the elaborate data model diagram showing the relationship between all the different tables and data layers in our model but happy to share it if anybody wants to. These are what you might call the use cases, or examples, starting with the simplest case, one parcel, one structure, one address – the address point ID, MP1 in this diagram, links the record to the point. We start out with a map of all the structures and a list of all the addresses, and the process of building the database consists of filling in the address point ID in the address record.



Master Address Database examples

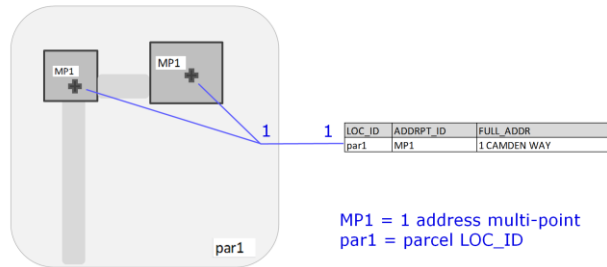
One ESL record doesn't match – field work is needed to find it. Utility and voter lists can be useful to validate such cases as well.

SOURCETYPE *	FULL_NUM *	FULL_STR *
WG	90	Blueberry Ln
L3	90	BLUEBERRY LN
ESL	94	BLUEBERRY LN
WG	94	Blueberry Ln
L3	94	BLUEBERRY LN
ESL	97	BLUEBERRY LN
WG	97	Blueberry Ln
ESL	100	BLUEBERRY LN
WG	106	Blueberry Ln
L3	106	BLUEBERRY LN
ESL	107	BLUEBERRY LN
WG	107	Blueberry Ln
L3	107	BLUEBERRY LN
ESL	114	BLUEBERRY LN



Here's a few examples of the field work – in this case a single blue dot represents the linear geocode of an address that doesn't match a structure. The address, 100 Blueberry Lane, doesn't look like there's room for it between 94 and 106, but the number is in the ESL, so we have to look for it, even though it looks like an error.

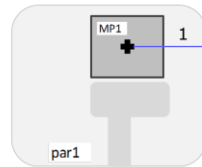
**Most common case –
several structures, one address (not a “site”)**



In aerial view, these structures on a residential lot could be a house, a barn, a garage, an in-law apartment. For public safety purposes, we can't assume we know which is the "primary" structure or which ones have land-lines.

This is most likely a house and a garage on a single family lot. both structures are linked to the address, because you can't always tell which is which, or maybe there's an in-law apartment or some other address that you don't know about. The point is that you shouldn't try to guess which is the primary structure, so we have a 2-part multi-point with ID MP1 again, linked to an address record.

Also very common – One building with many units



LOC_ID	ADDRPT_ID	FULL_ADDR
par1	MP1	104 NORTON AVENUE #10
par1	MP1	104 NORTON AVENUE #11
par1	MP1	104 NORTON AVENUE #12
par1	MP1	104 NORTON AVENUE #13
par1	MP1	104 NORTON AVENUE #14
par1	MP1	104 NORTON AVENUE #15
par1	MP1	104 NORTON AVENUE #16
par1	MP1	104 NORTON AVENUE #17
par1	MP1	104 NORTON AVENUE #18
par1	MP1	104 NORTON AVENUE #19

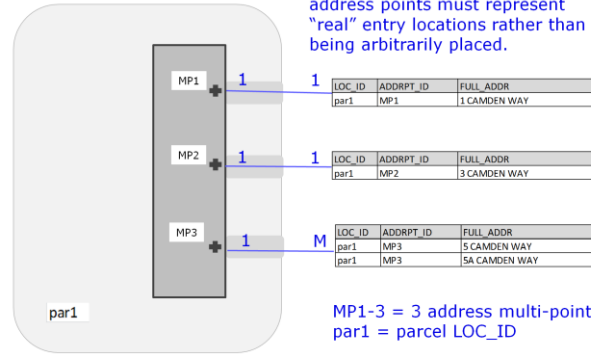
For this apartment building, a single address point is linked to multiple addresses, rather than being represented by “stacked” points. This relational approach makes editing and data management, especially in more complex cases covered later, much easier.

this is also very common, a multi-family dwelling. Again, we prefer managing a single point rather than a stack. so the address point ID gets repeated in every address record.

Higher geographic precision – one building, three entries, four numbered addresses



Here there are multiple addresses for a single structure. The address points must represent “real” entry locations rather than being arbitrarily placed.

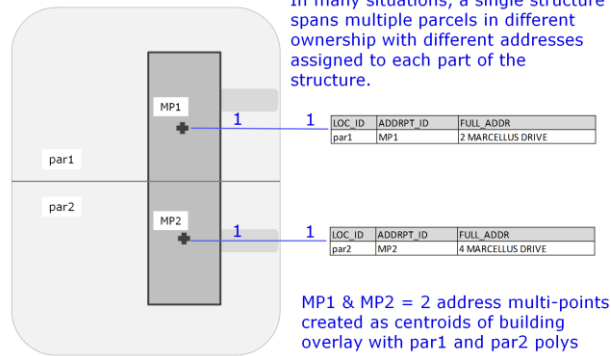


Now we're getting into more detail - this is a row house where you have separate numbered addresses, so the single address point has been replaced by 3 address points whose locations are collected in the field. each one has its own address point ID and is individually linked to an address record.

Other addresses appear as linear geocodes to indicated need for field verification. Address records not linked to structure will get addrpt_id for building interior points from fieldwork.

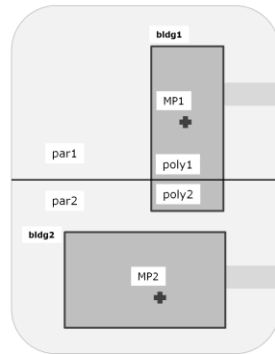


Two parcels, one building, two numbered addresses



if a building spans a parcel boundary and you have two different addresses then we automatically generate two points, again each is a distinct feature.

Two parcels split one building - when not to retain address point for both parts



MP1 = address multi-point created as centroid of building split by parcel boundary, additional multi-point in poly2 discarded if:
area poly2 < 15% area bldg1
or
area poly2 < 150 s.f.
or
area bldg2 > area poly2 *and*
area poly2 < 1000 s.f. *and*
area poly2 < 50% area bldg1

but not if:
no bldg2 and par2 has valid address
or year-built or building-value

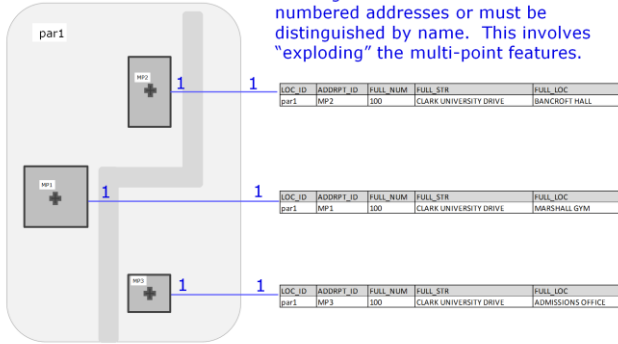
in all cases, perform manual review if
poly2 receives address point based
on above conditions

we don't always generate another point, if the building is just hanging over the boundary – these are the rules that we use to determine if two points are needed...but in the interests of time I won't go into that level of detail.

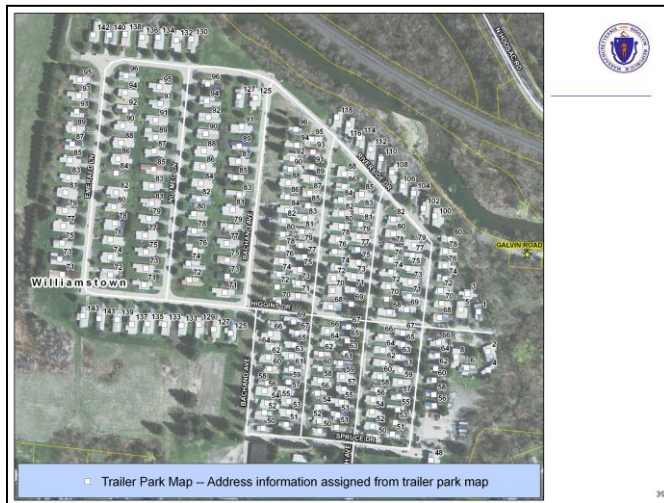
Campus – a “site” with many structures, each must be named



Separate address points are created when buildings on a site have different numbered addresses or must be distinguished by name. This involves “exploding” the multi-point features.

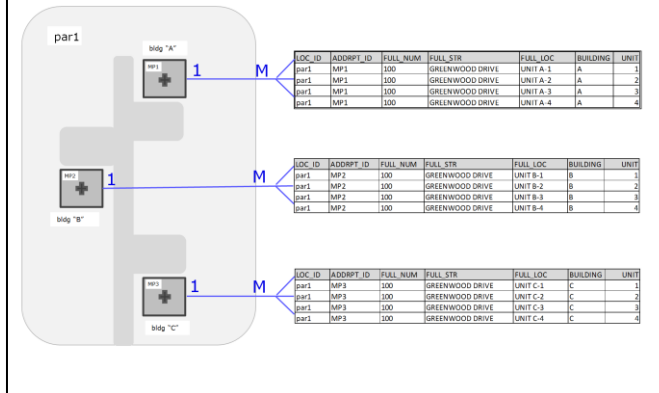


This is what we call a site, which is a polygon designated as an area within which individual buildings must be tagged for response purposes. So this is a mock campus, you have here the admissions building, the gym, bancroft hall, each addressed by name. You may aggregate parcels in common ownership to create sites, and you may in the field decide that you don't need to tag every structure, but that's the general idea..



We worked with Williamstown MA over a year ago on address assignment

Condo complex – a “site” with many structures, each with multiple units



This is very common and the most problematic of the site configurations – many units spread across many buildings – this site could be 40 acres in size and have dozens of buildings. In the field, the unit records are displayed in the field data capture app and have to be selected and assigned to the right building. The idea though is just the same as all the other cases, the address point ID is transferred by the field data capture app to the appropriate address records...

Master Address Database examples (designated site, before fieldwork)




"WESTWOOD
GLEN" record in tax
list is represented as
a multi-point.

Individual address
records are not
linked to any point
location and so
appear as linear
geocodes (light
blue).

SOURCETYPE *	FULL_NUM *	FULL_STR *
L3	<Null>	WESTWOOD GLEN
MSAG	1-161	WESTWOOD GLEN RD
ESL	153	WESTWOOD GLEN RD
ESL	154	WESTWOOD GLEN RD
ESL	159	WESTWOOD GLEN RD



Here's a condo complex, Westwood Glen in Westwood, MA, where you have seven big condo buildings and a total of 161 units to be assigned. I've shown just a few of the unit records. Here you see in the top record the assessor lists the address as just Westwood Glen; the MSAG shows Westwood Glen Road, which we have mapped so we can geocode the individual address numbers along that little stretch of Westwood Glen Road. You can see we've designated this as a site, which we know we have to visit. Each building has a single address point, and we want at a minimum to assign the right addresses to each building.




Master Address Database (after fieldwork)

After fieldwork, groups of records are associated with one or more points per building. Linear geocodes not needed.

SOURCETYPE *	FULL_NUM *	FULL_STR *
L3	<Null>	WESTWOOD GLEN
MSAG	1-161	WESTWOOD GLEN RD
ESL	153	WESTWOOD GLEN RD
ESL	154	WESTWOOD GLEN RD
ESL	159	WESTWOOD GLEN RD

Addrpt_id links the records in the MAF to the address point features



43

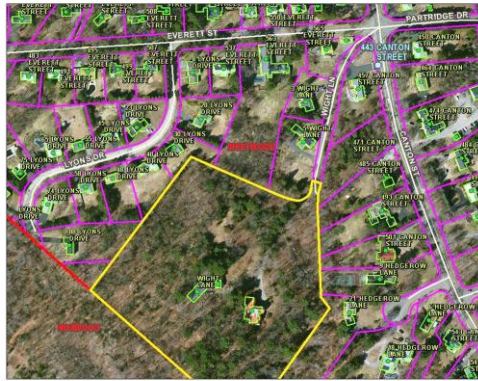
Here's how it is after field work – the addresses now are linked to points actually representing sections of the buildings with separate entrances, so the three address records would now be linked via address point ID to the point I've indicated...

Master Address Database examples – research required for some addresses



The parcel highlighted in orange has no number on Wight Lane in the assessors database. In fact, the address appears as 443 Canton Street in the ESL and other sources. Originally, Wight Lane was just a driveway for the house and not developed.

The addrpt_id for the two structures is added to the 443 Canton Street records.



44

Master Address Database – research required for some addresses



The pool house and landscaping on Creston Ave parcel in Westwood (highlighted in orange) appear to be associated with 17 Vincent Road in neighboring town.

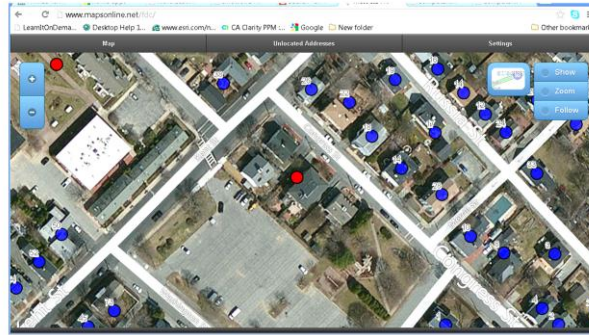
Since this parcel has a structure, we need to investigate addressing.



45



Tablet app to review addresses in the field



46



How will we find out about new addresses?

- **Right now, we get updates to the ESL and identify new and changed records, but land-lines, the source of data for the ESL, are going away**
- **we are trying to link into local workflows for address creation and maintenance**
- **RPA's did survey to identify "addressing authority" in each community**
- **Other potential sources of address data in real-time are:**
 - utilities, may capture location, limited number of sources, location will be electric drop, not always at building
 - ZIP plus four, monthly updates from PO



How will we locate new addresses?

- **In general, this will involve field work with GPS and some kind of mobile data collection device, similar to how we plan to clean up the draft point data**
- **In VT – 2 people drive around and do this work for the entire state**
- **In most states, counties maintain this information**
- **Best practice is to locate new address points inside structure outlines as they will appear on the next orthophoto, but that has practical constraints**
 - estimating exact location is hard, especially if you can't access building site
 - equipment does exist to collect distance and bearing to a target in the field along with GPS location of the observer and record that in a mobile device



What's next?

- **Version 2 of mobile app under development and will be complete by April of this year**
- **Datasets of points for field review are being prepared for each town**
 - locations needing numbered address
 - locations needing more address detail (site)
 - addresses we cannot locate at all
 - addresses we need to locate more precisely
- **We will be ready to provide data and support communities working in the field starting this spring**
- **We are merging local datasets into the statewide database where they can add value**
 - planimetrics (we will compare)
 - point datasets (ideally, conforming to our data model)

49



Feedback, Comments, Questions

We want to hear from you:

- **Do you have questions or concerns about this project?**
- **Are you trying to solve the same problem at the local level?**
- **What can we learn from your efforts and experience?**
- **The GIS data address location database will be an incredibly valuable resource for municipalities**
 - public safety, permitting, notifications, schools etc.

How are you currently using address data?

Most important, we want to engage you in this effort:

- **We are already working with many cities and towns to review the street map and the emergency service zones**
- **We want to work with communities to make the master address database as complete and accurate as it can be by doing field work**
- **For new addresses – send us email at Notify911Address@state.ma.us**
 - For new subdivisions we will get NAVTEQ to map the new streets and ranges
 - For all new addresses we will add that location to be field verified

50